Tạp chí Tin học và Điều khiển học, T.23, S.4 (2007), 356–363

THUẬT TOÁN GIẤU TIN HỖN HỢP

NGUYỄN NGỌC HÀ

Bưu điện Hải Phòng

Abstract. This paper presents an image data hiding algorithm composing two models of frequency and image spaces. In comparison of well-known algorithms this algorithm can support high capacity for each image block, robustness and cryptographic security. By dicrete cosine transfer DCT, we transfer from image spaces to frequency space, using EO algorithm to take middle frequency space, make 0, 1 matrix to hidding data , using Cheng-pan-Tseng algorithm to hidding , using invert dicrete cosine transfer IDCT to transfer image space, so the image has hidding data. When take data, using dicrete Cosine transfer DCT to transfer from image space to frequency space, and EO algorithm and Chang-Pan-Tseng algorithm to take hidding data.

Tóm tắt. Bài báo trình bày một thuật toán giấu tin trên cơ sở phối hợp hai mô hình dựa trên miền tần số và dựa trên không gian ảnh. So với các thuật toán đã biết, thuật toán này đảm bảo đồng thời được các tiêu chuẩn giấu được nhiều thông tin trong mỗi khối ảnh, bền vững với một số phép biến đổi và có độ bảo mật cao.Thông qua phép biến đổi Cosine rời rạc DCT, chúng ta đã chuyển được từ miền giá trị của ảnh (miền ảnh) sang miền tần số của ảnh (miền tần số), sau đó sử dụng thuật toán chẵn lẻ EO để trích miền tần số giữa của ảnh, tạo ra một khối ma trận nhị phân gồm các phần tử 0 và 1, để giấu dữ liệu vào miền này, sử dụng thuật toán Cheng- Pan- Tseng [5] để giấu dữ liệu trong miền chẵn lẻ, sau đó sử dụng phép biến đổi ngược chẵn lé IEO để biến đổi lại miền tần số, và sử dụng phép biến đổi ngược Cosine rời rạc IDCT để chuyển đổi về miền giá trị của ảnh. Như vậy ảnh sau khi biến đổi đã được giấu dữ liệu. Việc trích dữ liệu chỉ việc làm ngược lại , thông qua phép biến đổi Cosin rời rạc DCT chuyển qua miền tần số, và thuật toán chẵn lẻ EO để trích ra miền tần số giữa được ma trận nhị phân, sử dụng phép trích dữ liệu của Cheng-Pan-Tseng [5] để lấy dữ lệu.

1. BÀI TOÁN GIẤU TIN VÀ CÁC GIẢI PHÁP

Cho ảnh F và dữ liệu D thể hiện dưới dạng dãy bit. Yêu cầu giấu dữ liệu D trong ảnh F đảm bảo các tính chất: 1) Ảnh đích F' chứa dữ liệu D không sai khác nhiều so với ảnh nguồn F, chí ít là bằng mắt thường không thể cảm nhận được sự sai khác. 2) Ảnh đích F' bền vững đối với một số phép biến đổi ảnh. 3) Ảnh đích F' có thể chứa một dung lượng lớn dữ liệu D. 4) Có thể trích lại chính xác lượng tin D từ F'. 5) Các thủ tục giấu và trích tin hoạt động nhanh. 6) Đối phương khó dò tìm phát hiện dữ liệu D.

Có hai phương thức tiếp cận chủ yếu cho các thủ tục giấu tin trong ảnh: dựa trên không gian ảnh và dựa trên miền tần số [3, 5, 6]. Các thuật toán dựa trên không gian ảnh thao tác trực tiếp trên các điểm ảnh, các thuật toán dựa trên miền tần số coi mỗi dãy giá trị biểu diễn các điểm ảnh như một dãy giả ngẫu nhiên thể hiện tần số hoặc biên độ quan sát được trong một tiến trình giả định, thuật toán biến đổi dãy tần số này thông qua các phép biến đổi toán-lý như Fourie, cosin rời rạc hoặc sóng nhỏ.

Để hình thức hóa trong trình bày, bài báo tạm phân loại ảnh như sau: các ảnh nhị phân chỉ có hai điểm màu đen (0) và trắng (1), các ảnh nhiều màu (trên hai màu, gọi chung là ảnh màu bao gồm cả ảnh đa mức xám), mỗi màu được mã số bằng một số nguyên. Các ảnh được chia thành các khối kích thước $m \times n$. Các thuật toán đều thao tác trên các khối ảnh theo nghĩa: thuật toán $T(d, B)$ giấu lượng tin d vào khối ảnh B, thuật toán $IT(B, d)$ - trích lượng tin d từ khối ảnh chứa tin B. Các tham biến khác như khóa, các ma trận phụ trợ coi như cho trước. Để giấu lượng tin D trên toàn ảnh F ta chia D thành các đoạn $d_1, d_2, ..., d_k$ và chia ảnh F thành các khối rồi giấu mỗi đoạn tin d_i vào một khối ảnh $B; i = 1, 2, ..., k$. Để trích lượng tin D trên toàn ảnh, thoạt tiên ta trích các đoạn tin d_i từ mỗi khối ảnh chứa tin, sau đó ghép các đoạn tin này để thu được $D = (d_1, d_2, ..., d_k)$. Các thuật toán trích tin đề cập trong bài đều không đòi hỏi ảnh nguồn.

1.1. Biến đổi trên không gian ảnh

Các thuật toán giấu tin theo tiếp cận biến đổi trên không gian ảnh hoạt động theo sơ đồ chung mô tả dưới đây.

Quy trình giấu lượng tin d vào một khối ảnh màu B

Bước 1. Từ khối ảnh màu B trích ra một khối nhị phân G theo phép biến đổi: mỗi điểm màu sinh ra một bít 0/1: $EO(B, G)$.

Bước 2. Giấu lượng tin d vào khối ảnh nhị phân $G : DH(d, G)$.

Bước 3. Trả lại khối ảnh nhị phân G về khối ảnh màu $B : IEO(G, B)$.

Quy trình trích tin từ khối ảnh chứa tin

Bước 1. Từ khối ảnh chứa tin B trích ra một khối nhị phân $G : EO(B, G)$.

Bước 2. Trích lượng tin d từ khối nhị phân $G : IDH(G, d)$.

1.2. Biến đổi trên miền tần số

Các thuật toán giấu tin theo tiếp cận biến đổi trên miền tần số hoạt động theo sơ đồ chung mô tả dưới đây.

Quy trình giấu lượng tin d vào một khối ảnh màu B

Bước 1. Biến đổi khối ảnh màu B thành ma trận số $M : T(B, M)$.

Bước 2. Giấu lượng tin d vào ma trận $M : WM(d, M)$.

Bước 3. Biến đổi ngược M về $B : IT(M, B)$.

Quy trình trích tin từ khối ảnh chứa tin

Bước 1. Biến đổi khối ảnh màu chứa tin B thành mạ trận số $M : T(B, M)$.

Bước 2. Trích lượng tin d từ $M : IWM(M, d)$.

1.3. Biến đổi DCT

Phép biến đổi cosin rời rạc (Discrete Cosin Transform - DCT) lần đầu tiên được Ahmed và đồng nghiệp vận dụng vào năm 1974 [1, 2, 3, 5].

Phép biến đổi thuận DCT, cho ma trận bậc N với các chỉ số biến đổi từ 0 đến $N-1$ được định nghĩa như sau. Ký hiệu ma trận đầu vào là X, ma trận đầu ra là I, ta có$DCT(X, I)$

$$I[u, v] = \xi(u)\xi(v) \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} X[k, l] \cos(\frac{(2k+1)u\pi}{2N}) \cos(\frac{(2l+1)v\pi}{2N}),$$

các số thực $I[u, v]$ được gọi là hệ số DCT.

Biến đổi ngược $IDCT(I, X)$, được định nghĩa như sau

$$X[k,l] = \sum_{k=0}^{N-1}\sum_{v=0}^{N-1} \xi(u)\xi(v)I[u,v]\cos(\frac{(2k+1)u\pi}{2N})\cos(\frac{(2l+1)v\pi}{2N}),$$

trong đó $0 \le k, l, u, v \le N-1$ và

$$\xi(t) = \begin{cases} \frac{1}{\sqrt{N}} & \text{nếu } t = 0 \\ \sqrt{2/N} & \text{nếu } 1 \le t \le N-1. \end{cases}$$

Các thuật toán DCT và IDCT được cài đặt với độ phức tạp tính toán $O(N^2 \log N)$, trong đó N là bậc của khối ảnh, log được tính theo cơ số 2. Các hệ số DCT chứa thông tin về mật độ phân bố tần số không gian của thông tin trong khối. Khối hệ số DCT, I có thể chia thành 3 miền, miền tần số thấp, miền tần số giữa và miền tần số cao. Các thông tin trong miền tần số cao thường không mang tính tri giác cao. Miền tần số thấp cũng khó được sử dụng vì với một sự thay đổi dù nhỏ trong miền này cũng ảnh hưởng đến chất lượng tri giác của ảnh. Vì vậy, miền tần số ở giữa thường hay được sử dụng để giấu tin và cũng cho kết quả tốt nhất.

2. BẤT BIẾN

Bất biến là một mệnh đề $P(B,d)$ phát biểu trên khối ảnh B và dãy dữ liệu d như sau.

Cần giấu dữ liệu d vào khối ảnh B. (i) Nếu $P(B,d)$ thì coi như đã giấu d vào B. (ii) Nếu $notP(B,d)$ thì sửa B để thu được $P(B,d)$.

Ví dụ 1. (Thuật toán Wu_Lee [4]) B là khối ảnh nhị phân, d là dãy bít dữ liệu thể hiện như một số nguyên không âm, K là một khóa dạng khối nhị phân, W là một khóa trọng số chứa ít nhất một lần xuất hiện của các số $1, 2, ..., p-1, p = 2^r$, $0 \le d < p$. Các ma trận B, K và W cùng bậc. Kí hiệu $\oplus$ là phép toán cộng loại trừ (XOR) theo bit tương ứng của hai khối nhị phân cùng bậc, $\otimes$ là phép toán nhân các phần tử tương ứng của hai ma trận cùng bậc. Ta có thể mô tả bất biến của thuật toán giấu dãy bít d vào khối ảnh B có sử dụng khóa K và ma trận trọng số W như sau: $SUM((B \oplus K) \otimes W)\%p = d$.

Ví dụ 2. B là ma trận số được biến đổi DCT từ khối ảnh màu cho trước, d là một bit dữ liệu cần giấu trong B, khi đó một trong các bất biến có thể mô tả như sau [2, 3, 5]: tồn tại hai phần tử $B[i,j]$ và $B[p,q]$ trong B để: $((v \ge c) \wedge (d = 1)) \vee ((v < c) \wedge (d = 0))$,

$v = ||B[i,j]| - |B[p,q]||$, c là một số nguyên dương tùy chọn thích hợp.

Dễ thấy, nếu $v < c$ và $d = 1$ thì có thể sửa một trong hai hệ số $B[i,j]$ hoặc $B[p,q]$ để thu được bất biến $(v \ge c) \wedge (d = 1)$. Tương tự có thể xét cho trường hợp $v \ge c$ và $d = 0$.

3. ĐỘ NHÚNG TIN

Gọi P là lớp các thuật toán giấu tin thỏa các điều kiện sau đây:

- Ảnh nguồn được chia thành các khối kích thước $m \times n$.
- Mỗi khối ảnh giấu được r bit dữ liệu.
- Để giấu r bit dữ liệu như trên, thuật toán sửa tối đa k phần tử trong khối.

Đặt $t = r/k$ và gọi đại lượng này là tỷ lệ *nhúng/sửa*. Khi đó độ nhúng tin α được tính theo công thức $\alpha = r/(kmn) = (r/k)/mn$.

Trong hệ thức trên, đại lượng r/k cho biết tỷ lệ giữa lượng tin giấu được trong một khối và số điểm ảnh bị thay đổi trong khối. Hai thuật toán xử lý cùng một khối ảnh, tức là

cùng một diện tích $m \times n$ tính bằng pixel ảnh, thuật toán nào có tỷ lệ *nhúng/sửa* lớn hơn sẽ tốt hơn. Như vậy, độ nhúng tin là mọt trong những chỉ số đánh giá hiệu quả của các thuật toán giấu tin. Độ nhúng tin càng lớn thì thuật toán càng tỏ ra có hiệu quả. Ví dụ, xét thuật toán Wu_Lee với kích thước khối là 4×4 ($m = n = 4$), mỗi khối có thể giấu được tối đa 1 bit dữ liệu ($r = 1$) với điều kiện sửa tối đa 1 phần tử ($k = 1$) ([5]). Ta tính được $\alpha_1 = 1/(1 \times 4 \times 4) = 1/16$. Giả sử thuật toán DH sẽ trình bày dưới đây cũng chọn kích thước khối là 4×4, mỗi khối sẽ giấu tối đa 3 bit dữ liệu với điều kiện sửa tối đa 2 phần tử trong khối. Ta tính được $\alpha_2 = 3/(2 \times 4 \times 4) = 3/32$. Hệ thức $\alpha_2 > \alpha_1$ cho ta biết, trên cùng một khối có diện tích là $4 \times 4 = 16$ pixel ảnh, thuật toán thứ nhất có tỷ lệ nhúng/sửa là 1/1 - giấu 1 bit dữ liệu trên cơ sở sửa một điểm ảnh trong khối, trong khi thuật toán thứ hai có tỷ lệ nhúng/sửa là 3/2 - giấu 3 bit dữ liệu trên cơ sở sửa hai điểm ảnh trong khối.

Dĩ nhiên, như đã trình bày, để đánh giá đầy đủ hiệu quả của một thuật toán giấu tin, ngoài độ nhúng tin, ta còn phải xét các yếu tố khác như độ bảo mật, độ an toàn hay tính bền vững trước các phép tấn công (các phép biến đổi ảnh đích), tốc độ nhúng và trích tin...

4. THUẬT TOÁN DH – GIẤU TIN VÀO KHỐI ẢNH ĐEN TRẮNG

Các thuật toán giấu tin trong ảnh nhị phân tạo thành một lớp cơ sở để xây dựng các thuật toán giấu tin trong ảnh màu. Chính vì vậy mà việc tập trung nỗ lực nhằm hoàn thiện lớp cơ sở này là có ý nghĩa.

Phiên bản đầu tiên của thuật toán do nhóm nghiên cứu Yu-Yuan Chen, Hsiang-Kuang Pan và Yu-Chee Tseng của Đại học Quốc gia Chung-Li, Taiwan công bố vào năm 1998 [4]. Những ý tưởng chính của thuật toán là:

- Sử dụng một ma trận trọng số W nhằm gia tăng tỉ lệ tin giấu.

- Sửa mỗi khối không quá 2 bit nhưng có thể giấu $r \geq 2$ bít thông tin với $r = \lfloor \log(mn) \rfloor$, trong đó $\lfloor x \rfloor$ là kí hiệu phần nguyên của x.

Ta kí hiệu $DH(d, B)$ là thuật toán giấu dãy r bit d vào khối ảnh đen trắng B và $IDH(B, d)$ là thuật toán trích dãy bit d từ khối ảnh chứa tin B [4].

5. THUẬT TOÁN HỖN HỢP

Thuật toán DH cho phép giấu nhiều bit dữ liệu trong mỗi khối, tuy nhiên độ bền vững không cao, trái lại, nếu sử dụng phép biến đổi DCT và giấu tin vào vùng tần số giữa thì có thể đạt được độ bền vững cao, tuy nhiên chỉ giấu được ít thông tin, thông thường là một bít trong mỗi khối. Vì các lý do đó, mỗi lớp thuật toán được khai thác trong một lĩnh vực khác nhau. Thuật toán DH thích hợp với các trường hợp cần giấu nhiều dữ liệu và thời gian tồn tại trên đường truyền là rất ngắn, không có mục đích phát tán rộng rãi, thí dụ, một đề thi đã mã hóa, một bản hợp đồng đã được ký bằng chữ ký số. Thuật toán sử dụng phép biến đổi DCT thích hợp với các trường hợp bảo vệ bản quyền cho các đối tượng (ảnh) để công khai và lâu dài trên mạng máy tính, hoặc các trường hợp cần bảo vệ đặc biệt bằng cách nhúng một thủy vân vào các đối tượng đó, thí dụ, một văn bằng cần gửi và trao đổi trên mạng, một bức ảnh dự triển lãm điện tử trên mạng hoặc một khóa cần gửi đến các hội đồng thi để mở đề thi [3, 5]...

Một nhận xét tự nhiên là nếu kết hợp hai kỹ thuật nói trên một cách khoa học thì có thể nhận được một thuật toán đáp ứng đồng thời hai yêu cầu xem như trái ngược nhau: vừa

giấu được nhiều dữ liệu vừa có độ bền vững cao. Đây là một trong những kết quả chủ yếu của bài báo. Thuật toán được triển khai có tên là DHT.

5.1. Thuật toán DHT: giấu tin vào khối ảnh màu

Algorithm DHT;
Function: Giấu số nhị phân r bit d vào khối ảnh B
Input
- Khối ảnh nguồn B bậc m
- Số nhị phân d gồm r bit $d = (d_1, d_2, ..., d_r)$
- Khóa nhị phân K bậc n (cho trước)
- Ma trận trọng số W bậc n (cho trước)

Output
- Khối ảnh B chứa d

Format $DHT(d, B)$
Method

1. Thực hiện phép biến đổi DCT trên khối ảnh B để thu được ma trận C bậc m: $DCT(B, C)$;
2. Tạo ảnh nhị phân E bậc n từ miền tần số giữa của C bằng thủ tục chẵn lẻ EO : $EO(C, E)$;
3. Giấu số d vào E theo thuật toán DH : $DH(d, E)$;
4. Trả lại các bit từ E về C bằng thủ tục ngược với thủ tục chẵn lẻ$IEO(E, C)$;
5. Gọi thủ tục $IDCT$ để biến đổi ngược C về B và trả kết quả: $IDCT(C, B)$;

EndDHT.

Thuật toán có độ phức tạp $O(m^2 \log(m))$ vì các bước 1 và 5 có độ phức tạp cao nhất thực hiện các phép biến đổi DCT và IDCT trên các ma trận bậc m đòi hỏi thời gian tính toán $O(m^2 \log(m))$.

5.2. Thuật toán chọn miền tần số giữa EO

Trong bước 2 cuả thuật toán DHT ta dựa vào vùng giữa của ma trận C bậc m để tạo ra một ảnh nhị phân E bậc n bằng kỹ thuật chẵn lẻ. Thuật toán này có tên là EO và hoạt động như sau.

Vì bậc của ma trận C là m, trong khi bậc của ma trận E là $n < m$, nên để trích miền tần số giữa của C ta có thể dùng một mặt nạ (nhị phân) M, trong đó nếu giá trị $M[u, v] = 1$ thì ta lấy phần tử $C[u, v]$ tương ứng, ngược lại, khi $M[u, v] = 0$ thì ta bỏ phần tử đó. Như vậy mặt nạ M quy định vùng các tần số giữa của ma trận C. Đương nhiên ta phải xây dựng ma trận M sao cho $SUM(M) \geq n^2$, tức là số lương bit 1 trong M phải không nhỏ hơn số lượng phần tử trong ma trận E.

Thủ tục xác định tính chẵn lẻ đơn giản như sau. Gọi $C[u, v]$ là phần tử được chọn để phát sinh trị cho phần tử $E[i, j]$ của ma trận E. Như trên đã nói, $C[u, v]$ được chọn khi và chỉ khi $M[u, v] = 1$. Nếu phần nguyên của trị tuyệt đối của $C[u, v]$ là một số chẵn thì $E[i, j]$ nhận trị 0, ngược lại, khi $|C[u, v]|$ là một số lẻ thì gán $E[i, j] := 1$. Cụ thể là

$$E[i, j] := INT(abs(C[u, v])).$$

Hệ thức trên tương đương với việc lấy bit hàng đơn vị trong dạng biểu diễn nhị phân của phần tử $C[u,v]$. Một yêu cầu nữa là phải chọn các giá trị m, n và r sao cho $n < m$, và $n^2 \geq 2^r - 1$.

Hệ thức thứ hai được suy từ yêu cầu ma trận trọng số W bậc n phải chứa đủ các giá trị $1, 2, ..., 2^r - 1$. Theo các kết quả thực nghiệm (triển khai tại Viện Công nghệ Thông tin và Trung tâm Tin học Bưu điện Hải Phòng) thì miền tần số giữa trong C chiếm khoảng 1/3 số lượng phần tử của C, tức là khoảng $(1/3)m^2$. Như vậy, với m cho trước, nên chọn $n \leq m/\sqrt{3}$. Ví dụ, với $m = 8$ thì có thể chọn $n = 4$, từ đó suy ra giá trị cần chọn của r phải thỏa hệ thức $n^2 \geq 2^r - 1$, chẳng hạn chọn, $r = 3$. Tóm lại, nếu biết bậc m của các khối ảnh nguồn thì các giá trị lớn nhất có thể có của n và r là như sau:

$$n = \lfloor m/\sqrt{3} \rfloor,\ r = \lfloor 2\log_2 n \rfloor = \lfloor 2\log_2(m/\sqrt{3}) \rfloor.$$

Sau khi xác định được n và r ta có thể tính được độ nhúng tin của thuật toán DHT là

$$\alpha = r/(km^2) = 2\log_2(m/\sqrt{3})/(2m^2) = \log_2(m/\sqrt{3})/m^2.$$

Trong công thức trên, $k = 2$ là số pixel ảnh bị sửa trong mỗi khối để có thể giấu được r bit dữ liệu.

Trong thuật toán EO mô tả dưới đây, ma trận mặt nạ M được cho trước như một tham số tổng thể. Các chỉ số duyệt các phần tử của ma trận C và mặt nạ M là u (chỉ số dòng) và v (chỉ số cột), $1 \leq u, v \leq m$. Điều kiện để xét phần tử $C[u,v]$ là $M[u,v] = 1$. Các chỉ số cho các phần tử của ma trận E là i (chỉ số dòng) và j (chỉ số cột), $1 \leq u, v \leq n$. Các biểu thức $i := (i\%n) + 1$ và $j := (j\%n) + 1$ điều khiển tuần hoàn cho các chỉ số i và j biến thiên trong khoảng $1, ..., n$. Khi điền đủ n^2 giá trị cho E thì dừng thuật toán.

Algorithm EO;

Function: Tạo ma trận nhị phân E bậc n từ ma trận C bậc m bằng kỹ thuật chẵn lẻ thông qua ma trận mặt nạ M.

Input

- Ma trận C bậc m
- Mặt nạ miền tần số giữa M bậc m (cho trước)

Output

- Ma trận nhị phân E bậc n

Format $EO(C, E)$

Method

1. Khởi trị $i := 1; j := 1;$
2. for $u := 1$ to m do
   for $v := 1$ to m do
   if $(M[u,v] = 1)$ then
   $E[i,j] := INT(abs(C[u,v]))\ \%\ 2;$
   $j := (j\%n) + 1;$
   if $(j = 1)$ then
   $i := (i\%n) + 1;$ //chuyển dòng
   if $(i = 1)$ then stop endif;
   endif;
   endif;
   endfor v;
   endfor u;

EndEO.

Thuật toán biến đổi ngược so với thuật toán EO là IEO. Thuật toán này trả lại các bit từ ma trận E về ma trận C và hoạt động tương tự như EO với những thao tác đảo chiều. Với số thực r trong ma trận C, nếu muốn đưa bit b vào r ta làm như sau:

- Tách r thành hai phần: phần nguyên i và phần thập phân p.
- Đặt bit b vào hàng đơn vị của i thu được i'.
- Ghép i' với p để thu được r'.

5.3. Thuật toán trích tin IDHT

Thuật toán $IDHT$ dưới đây thực hiện qui trình trích đoạn dữ liệu d gồm r bit từ khối B theo khóa nhị phân K bậc n, ma trận trọng số W bậc n. Quy trình này, về cơ bản là ngược với quy trình mô tả trong thuật toán DHT. Trong thuật toán có sử dụng mặt nạ M để lấy các phần tử trong miền tần số giữa của ma trận C nhận được từ phép biến đổi DCT trên khối ảnh B. Mặt nạ này coi như được cho trước dưới dạng tham số tổng thể.

Algorithm IDHT

Function: Trích dãy r bit dữ liệu d từ khối ảnh B

Input

- Khối ảnh B bậc m
- Khóa nhị phân K bậc n (cho trước)
- Mặt nạ miền tần số giữa M bậc m (cho trước)
- Ma trận trọng số W bậc n (cho trước)

Output

- Dãy dữ liệu $d = (d_1, d_2, ..., d_r)$

Format $IDHT(B, d)$

Method

1. Thực hiện phép biến đổi DCT trên khối ảnh B bậc m để thu được ma trận hệ số C cùng bậc m: $DCT(B, C)$;
2. Tạo ảnh nhị phân E bậc n từ ma trận C bậc m bằng thủ tục chẵn lẻ EO : $EO(C, E)$;
3. Trích tin từ E bằng thuật toán IDH và cho ra kết quả $IDH(E, d)$;

EndIDHT.

5.4. Các đặc trưng của họ các thuật toán DHT và IDHT

Định lý 5.1. *Với các giá trị m, n, r, K, W, M phù hợp, các thuật toán DHT và $IDHT$ là đúng đắn theo nghĩa sau:*

1) Nếu nhúng đoạn dữ liệu d gồm r bit vào khối ảnh B bằng thuật toán DHT thì có thể trích lại đúng đoạn dữ liệu d thông qua thuật toán $IDHT$.

2) Độ phức tạp của thuật toán DHT và $IDHT$ là $O(m^2 \log(m))$.

3) Độ an toàn của thuật toán DHT là $2^s C_s^p p! q^{s-p}$, với $s = mn$, $p = 2r - 1$, q là một số tự nhiên tùy ý.

Chứng minh

Thuật toán DHT nhúng đoạn tin d dài r bit vào khối tần số B thông qua lời gọi các thuật toán DCT, DH và $IDCT$. Thuật toán $IDHT$ trích lại đoạn tin d thông qua lời gọi các thuật toán DCT và IDH. Các thuật toán DCT và $IDCT$ đã được chứng minh là đúng đắn [3, 5]. Tính đúng của thuật toán DH và IDH đã được chứng minh trong [4], do đó thuật toán DHT và $IDHT$ là đúng đắn.

Trong các bước thực hiện của các thuật toán DHTvà $IDHT$, phép biến đổi DCTvà $IDCT$ trên các ma trận bậc m đòi hỏi độ phức tạp $O(m^2\log(m))$, $m > n$, do đó độ phức tạp của các thuật toán DHTvà $IDHT$ là $O(m^2\log(m))$.

Trong trường hợp xấu nhất, ta giả thiết độ an toàn của thuật toán DHT có cùng mức với độ an toàn của thuật toán DH. Trong thực tế, việc bố trí các bit trong ma trận mặt nạ M dùng để khởi trị cho ma trận nhị phân E bậc n thông qua ma trận phân bố tần số C bậc m (các thuật toán EO và IEO) sẽ nâng độ an toàn của thuật toán thêm một hệ số thể hiện tổ hợp phân bố n^2 giá trị 1 trong ma trận nhị phân bậc m, $C_{m^2}^{n^2}$.

Tuy nhiên việc bỏ nhân tử này là hợp lý vì đối phương có thể dự đoán rằng thuật toán sẽ chỉ quan tâm đến các phần tử nằm trong miền giữa của ma trận biến đổi tần số qua các phép biến đổi DCT.

6. THỬ NGHIỆM VÀ ĐÁNH GIÁ

Thuật toán được thử nghiệm trong môi trường Matlab trên máy PC với các ảnh Lena và các file dữ liệu dạng text dung lượng 20-50K. Ảnh đích thu được có sai khác với ảnh nguồn 1-2 bit trong một số khối 8×8. Số lượng khối sai khác chiếm khoảng 15% tổng số khối trong ảnh. Nếu giấu tin tại nhiều vị trí thì ảnh tỏ ra khá bền vững đối với các tấn công ngẫu nhiên. Với 5 vị trí giấu trở lên thì luôn luôn khôi phục lại được toàn văn.

Qua thử nghiệm, xin đưa ra một số đề xuất sau:

Thuật toán hỗn hợp có thể sử dụng để gửi các đoạn tin ngắn với thời gian tồn tại trên đường truyền không lâu. Hiệu quả của thuật toán được thể hiện tốt trong trường hợp phối hợp với các thủ tục bảo mật và xử lý khác như mã hóa công khai, nén dữ liệu.

TÀI LIỆU THAM KHẢO

[1] Vũ Ba Đình, Nguyễn Hồng Hải, Nguyễn Xuân Huy, Xây dựng phương pháp giấu tin bền vững trong cơ sở dữ liệu không gian, *Kỷ yếu Hội thảo Quốc gia: Một số vấn đề chọn lọc của Công nghệ Thông tin*, Thái Nguyên, 29-31/08/2003, NXB Khoa học Kỹ thuật, Hà Nội (83–88).

[2] Nguyễn Xuân Huy, Trần Quốc Dũng, Một thuật toán thủy vân ảnh trên miền DTC, *Tạp chí Bưu chính Viễn thông: Các công trình nghiên cứu - triển khai viễn thông và CNTT* **10** (3) (2003) 89–94.

[3] Chun-Shien Lu, Multimedia Security: Steganography and Digital Watermarking Techniques for Protection of Intellectual Property, Idea Group Publishing, ISBN, 2004.

[4] M. Y. Wu and J. H. Lee, A novel data embedding method for two-color facsimile images, *Proceedings of International Symposium on Multimedia Information Processing*, Chung-Li, Taiwan, R.O.C., 1989.

[5] Yu-Yuan Chen, Hsiang-Kuang Pan, and Yu-Chee Tseng, "A secure data hiding scheme for two-color images", Department of Computer Science and Information Engnineering, National Central University, Chung-Li, 32054 Taiwan. (1998)

[6] www.watermarkingworld.org

Nhận bài ngày 15 - 7 - 2007
Nhận lại sau sửa ngày 17 - 8 -2007